สำนักงานวิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

บทที่ 2

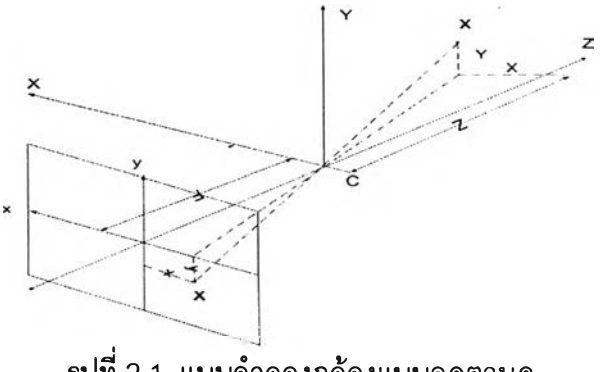
การทำภาพให้เสถียรทางหมุน

การแก้ปัญหาสัญญาณภาพวีดิทัศน์ที่ไม่เสถียรนั้นจำเป็นต้องศึกษาลักษณะการเปลี่ยนแปลง และความสัมพันธ์ระหว่างสัญญาณภาพภาพที่อยู่ติดกัน ซึ่งสามารถอธิบายได้หลายลักษณะ เช่น การเคลื่อนที่ของวัตถุในภาพ การเปลี่ยนแปลงความสว่างของภาพ หรือการเคลื่อนที่ของกล้อง ดังนั้นจึงต้องมีแบบจำลองเพื่อใช้อธิบายความสัมพันธ์ระหว่างการเปลี่ยนแปลงที่เกิดขึ้นจริงกับการเปลี่ยนแปลงที่เกิดขึ้นกับสัญญาณภาพ

2.1 แบบจำลองกล้อง (Camera Model)

แบบจำลองกล้อง คือแบบจำลองที่ใช้อธิบายการฉายวัตถุจากระนาบ 3 มิติหรือระนาบจริงลงสู่ระนาบ 2 มิติหรือระนาบภาพของกล้อง โดยแบบจำลองกล้องที่นิยมใช้กันอย่างแพร่หลายคือแบบจำลองกล้องแบบจุดตามด (Pinhole model) [4]

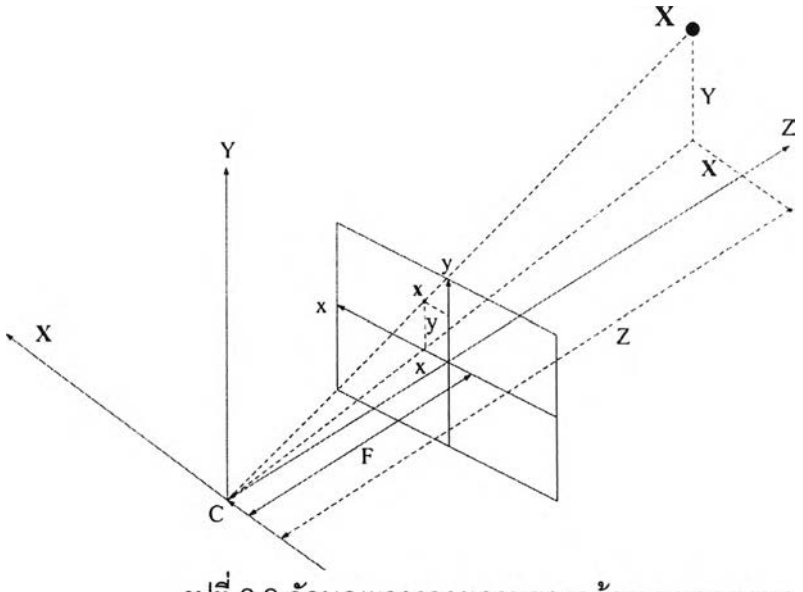
2.1.2 แบบจำลองกล้องแบบจุดตามด (Pinhole model)



รูปที่ 2.1 แบบจำลองกล้องแบบจุดตามด

รูปที่ 2.1 แสดงแบบจำลองกล้องแบบจุดตามด ในรูป F คือระยะโฟกัสของกล้อง C คือจุดศูนย์กลางโฟกัสหรือจุดศูนย์กลางของกล้อง ตำแหน่งของ x บนระนาบภาพเป็นตำแหน่งของวัตถุ X จากระนาบ 3 มิติที่แสดงบนระนาบภาพ ซึ่งเกิดจากการตัดกันของเส้นที่ลากเชื่อมระหว่าง X Cกับระนาบภาพ จากรูปที่ 2.1 ภาพที่ได้บนระนาบภาพ 2 มิติจะมีลักษณะผันกลับกับวัตถุจริงในระนาบ 3 มิติเนื่องจากระนาบภาพนั้นแสดงอยู่ด้านหลังของจุดศูนย์กลางโฟกัส ด้วยเหตุนี้ในกระบวนการทางวีดิทัศน์ส่วนใหญ่จึงกำหนดแบบจำลองให้แสดงระนาบภาพอยู่ด้านหน้าซึ่งก็คือ

อยู่ระหว่างจุดศูนย์กลางโฟกัสและตำแหน่งของวัตถุใน 3 มิติ ดังรูปที่ 2.2 เพื่อหลีกเลี่ยงปัญหาการผันกลับของระนาบภาพ และง่ายต่อการสังเกตประมวลผล



รูปที่ 2.2 ลักษณะการฉายภาพของกล้องแบบจุดตามด

ในรูปที่ 2.2 สมมติให้ระนาบ XY คือระนาบของวัตถุใน 3 มิติขนานกับระนาบภาพ ส่วนระบบพิกัด (X, Y, Z) ให้เป็นตามกฎมือขวาโดยกำหนดให้ทิศบวกของแกน Z คือทิศทางของภาพ ดังนั้นจากรูปที่ 2.2 ใช้คุณสมบัติสามเหลี่ยมคล้ายเขียนความสัมพันธ์ระหว่างระนาบภาพและระนาบวัตถุได้ดังนี้

$$\frac{x}{F}=\frac{X}{Z}, \qquad \frac{y}{F}=\frac{Y}{Z}$$
$$x=F\frac{X}{Z}, \qquad y=F\frac{Y}{Z} \tag{2.1}$$

F คือระยะโฟกัสของกล้อง

Z คือระยะจากวัตถุถึงจุดศูนย์กลางของกล้อง

x, y คือขนาดของภาพ

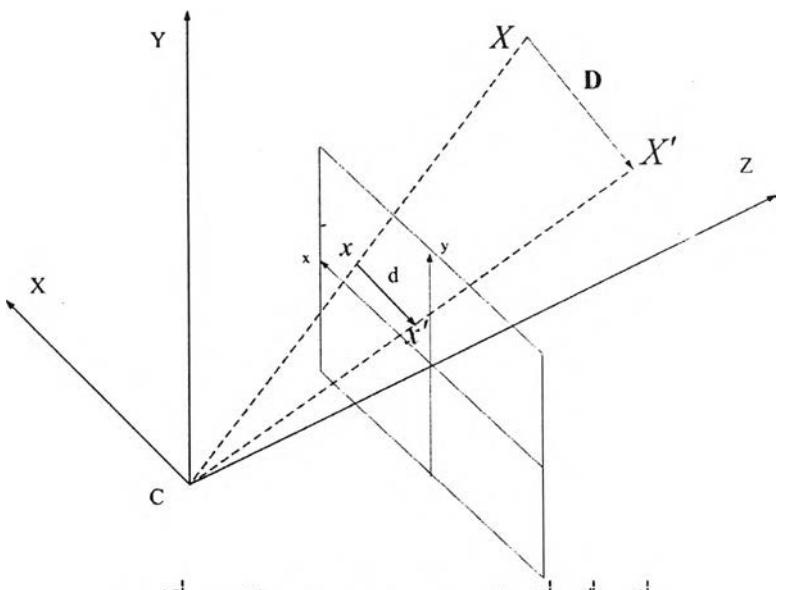
X, Y คือขนาดของวัตถุ

ความสัมพันธ์ระหว่างระนาบภาพและระนาบวัตถุนี้เรียกว่า เส้นโครงทัศนมิติ (Perspective projection) จุดที่น่าสังเกตของเส้นโครงทัศนมิติคือ เมื่อระยะห่างระหว่างวัตถุกับกล้องเพิ่มขึ้น ขนาดของภาพของวัตถุที่ได้ก็จะมีขนาดเล็กลงเช่นกัน ซึ่งลักษณะเช่นนี้สามารถอธิบายได้โดยความสัมพันธ์ผกผันระหว่างขนาด x และ y กับความลึก Z จากความสัมพันธ์ที่กล่าวมาแสดงให้เห็นว่าเมื่อวัตถุหรือกล้องเกิดการเคลื่อนที่ก็จะส่งผลให้ภาพที่ได้เกิดการเปลี่ยนแปลงด้วย จึง

สามารถนำความสัมพันธ์นี้ไปใช้หาการเปลี่ยนแปลงของระนาบภาพได้ ซึ่งลักษณะการเปลี่ยนแปลงหรือเคลื่อนที่ของระนาบภาพสามารถอธิบายได้โดยแบบจำลองการเคลื่อนที่

2.2 แบบจำลองการเคลื่อนที่ 2 มิติ (Two-dimensional motion model)

แบบจำลองการเคลื่อนที่ 2 มิติ คือแบบจำลองที่ใช้อธิบายความสัมพันธ์ระหว่างระนาบวัตถุกับระนาบภาพ นั้นคือเมื่อกล้องหรือวัตถุเกิดการเปลี่ยนแปลง ภาพของวัตถุนั้นก็จะเปลี่ยนแปลงด้วย จากรูปที่ 2.3 แสดงการฉายภาพวัตถุเคลื่อนที่ลงบนระนาบภาพ และแสดงความสัมพันธ์ระหว่างการเคลื่อนที่ใน 3 มิติและ 2 มิติ เมื่อวัตถุใน 3 มิติเคลื่อนที่จากจุด $\boldsymbol{X}=[X,Y,Z]^T$ ที่เวลา t_1 ไปที่จุด $\boldsymbol{X}'=[X',Y',Z']^T=[X+D_x,Y+D_y,Z+D_z]^T$ ที่เวลา $t_2=t_1+d_t$ ระนาบภาพจะเปลี่ยนแปลงจาก $\boldsymbol{x}=[x,y]^T$ ไปที่ $\boldsymbol{x}'=[x',y']^T=[x+d_x,y+d_y]^T$ เรียกการเลื่อนที่ ใน 3 มิติ $\boldsymbol{D}(\boldsymbol{X};t_1,t_2)=\boldsymbol{X}'-\boldsymbol{X}=[D_x,D_y,D_z]^T$ ว่าเวกเตอร์การเคลื่อนที่ 3 มิติที่จุด $\boldsymbol{X}$ และเรียกการเลื่อนที่ใน 2 มิติ $\boldsymbol{d}(\boldsymbol{x};t_1,t_2)=\boldsymbol{x}'-\boldsymbol{x}=[d_x,d_y]^T$ ว่าเวกเตอร์การเคลื่อนที่2 มิติที่จุด $\boldsymbol{x}$



รูปที่ 2.3 ลักษณะภาพฉายของวัตถุที่เคลื่อนที่

ลักษณะการเปลี่ยนแปลงของกล้องและวัตถุใน 3 มิติ นั้นอาจเกิดได้ทั้งแบบที่กล้องเคลื่อนที่ วัตถุอยู่กับที่ กล้องอยู่กับที่วัตถุเคลื่อนที่ หรือทั้งกล้องและวัตถุเคลื่อนที่ ซึ่งการเคลื่อนที่นั้นสามารถแยกได้เป็นการเลื่อนที่ และการหมุน แต่เพื่อให้ง่ายต่อการพิจารณาจึงกำหนดให้ส่วนใดส่วนหนึ่งอยู่กับที่ ในที่นี้กำหนดให้กล้องอยู่กับที่ แต่วัตถุเคลื่อนที่

ในกรณีทั่วไปความสัมพันธ์ของวัตถุก่อนและหลังการเคลื่อนที่สามารถเขียนเป็นสมการได้ดังนี้

$$X' = rX + T \tag{2.2}$$

$$\begin{bmatrix} X' \\ Y' \\ Z' \end{bmatrix} = \begin{bmatrix} r_1 & r_2 & r_3 \\ r_4 & r_5 & r_6 \\ r_7 & r_8 & r_9 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} + \begin{bmatrix} T_x \\ T_y \\ T_z \end{bmatrix} \tag{2.3}$$

จากสมการที่ 2.2 และ 2.3 X' คือตำแหน่งของวัตถุหลังการเคลื่อนที่ X คือตำแหน่งของวัตถุก่อนการเคลื่อนที่ เมทริกซ์ r คือเมทริกซ์การหมุนในแนวแกน x แกน y และแกน z ใช้แทนค่าอัตราการขยาย การหมุน การกวาด และการก้มเงยของกล้อง และเมทริกซ์ T คือการเคลื่อนที่ตามแนวแกน x แกน y และแกน z จากสมการที่ 2.1 สามารถเขียนความสัมพันธ์ระหว่างตำแหน่งของภาพก่อนและหลังการเคลื่อนที่ได้ดังนี้

$$\begin{aligned} x' &= F\frac{(r_1x + r_2y + r_3F)Z + T_xF}{(r_7x + r_8y + r_9F)Z + T_zF} \\ y' &= F\frac{(r_4x + r_5y + r_6F)Z + T_yF}{(r_7x + r_8y + r_9F)Z + T_zF} \end{aligned} \tag{2.4}$$

จากสมการถ้าทำการสเกลค่า T และ Z ด้วยตัวประกอบตัวเดียวกัน จะไม่ส่งผลต่อความสัมพันธ์ระหว่าง (x, y) และ (x', y') เพราะค่าของ T และ Z ขึ้นอยู่กับตัวประกอบมาตราส่วน (scaling factor) นั้นก็คือจะสามารถหาค่าทิศทางของ T_x T_y และ T_z ได้ แต่ไม่สามารถหาค่าความยาวได้

2.2.1 การแปลงเชิงภาพฉาย (Projective Mapping)

จากแบบจำลองการเคลื่อนที่ในสมการที่ 2.4 เมื่อภาพไม่มีการเคลื่อนที่ในแนวแกน z เรียกลักษณะการเปลี่ยนแปลงนี้ว่า การแปลงเชิงภาพฉาย ซึ่งเกิดการประมาณลดค่าตัวแปรในแกน z กำหนดให้ F มีค่าเป็น 1 และการลดเมทริกซ์ T ดังนั้นสามารถเขียนสมการที่ 2.4 ได้เป็น

$$\begin{aligned} x' &= \frac{a_0 + a_1x + a_2y}{1 + c_1x + c_2y} \\ y' &= \frac{b_0 + b_1x + b_2y}{1 + c_1x + c_2y} \end{aligned} \tag{2.5}$$

ดังนั้นค่าการเลื่อนที่ (displacement) ของการแปลงเชิงภาพฉายซึ่งคือค่าที่บอกการเปลี่ยนแปลงของภาพสามารถเขียนเป็นสมการได้ดังนี้

$$x' = x + d_x(x,y)$$
$$y' = y + d_y(x,y) \quad (2.6)$$

$$d_x(x,y) = \frac{a_0 + a_1 x + a_2 y}{1 + c_1 x + c_2 y} - x$$
$$d_y(x,y) = \frac{b_0 + b_1 x + b_2 y}{1 + c_1 x + c_2 y} - y \quad (2.7)$$

$d_x(x,y)$ คือค่าการเลื่อนที่ในแนวแกน x และ $d_y(x,y)$ คือการเคลื่อนที่ในแนวแกน y

2.2.2 การแปลงสัมพรรค (Affine Mapping)

การแปลงสัมพรรค คือการประมาณค่าการแปลงเชิงภาพฉาย โดยจะลดตัวแปรที่เกี่ยวข้องกับค่าเปลี่ยนแปลงความถี่ที่มีในการแปลงเชิงภาพฉายลง ซึ่งยังสามารถใช้อธิบายลักษณะการเปลี่ยนแปลงของภาพส่วนใหญ่ได้ ยกเว้นในกรณีที่เกิดปรากฏการณ์การเปลี่ยนแปลงความถี่ในภาพนั้นคือปรากฏการณ์ chirping effect และปรากฏการณ์ตีบแสง (converging effect) โดยปรากฏการณ์ chirping effect คือปรากฏการณ์ที่ค่าความถี่กำหนดทางระยะทาง (perceive spatial frequency) หรือค่าความถี่ในภาพมีค่าเพิ่มขึ้น เมื่อระยะห่างระหว่างกล้องกับวัตถุเพิ่มขึ้น ส่งผลให้ภาพมีลักษณะคล้ายโดนบีบเล็กลง ส่วนปรากฏการณ์ตีบแสง (converging effect) คือปรากฏการณ์การเคลื่อนตัวของเส้นขนานบนภาพเข้าใกล้จนอาจเกิดเป็นเส้นเดียว ดังนั้นสามารถเขียนค่าการเลื่อนที่ (displacement) ของการแปลงสัมพรรคได้ดังสมการต่อไปนี้

$$d_x(x,y) = a_0 + a_1 x + a_2 y$$
$$d_y(x,y) = b_0 + b_1 x + b_2 y \quad (2.8)$$

โดยการแปลงสัมพรรคสามารถหาได้จากความสัมพันธ์ของจุดลักษณ์(จุดพิกัดของจุดสนใจบนภาพ)ระหว่างลำดับภาพจำนวน 3 คู่จุดลักษณ์ ซึ่งค่าคู่พิกัดจุดลักษณ์นี้นำไปใช้ในการแก้สมการหาค่าตัวแปร a_0 a_1 a_2 b_0 b_1 และ b_2 ในสมการที่ 2.8 ซึ่งจะได้ค่าการเลื่อนที่ $d_x(x,y)$ และ $d_y(x,y)$ เพื่อนำไปใช้ชดเชยการเคลื่อนที่ของภาพ การหาความสัมพันธ์ระหว่างจุดลักษณ์ก็คือการจับคู่ของจุดลักษณ์ที่เปลี่ยน หรือเลื่อนที่ไปในแต่ล่ะภาพ หรือก็คือการติดตามจุดลักษณ์ไปในทุกลำดับภาพ

2.3. การติดตามจุดลักษณ์ (Feature tracker)

2.3.1 การตรวจหาจุดลักษณ์ (Feature detection)

การตรวจหาจุดลักษณ์ คือการเลือกจุดพิกัดจากภาพเพื่อใช้ในการสังเกตหรือติดตามการเปลี่ยนแปลง ซึ่งต้องเป็นจุดที่เห็นได้ชัดในภาพ เช่น จุดมุม จุดตัดเส้น หรือจุดต่อ ในที่นี้เลือกวิธีการตรวจหาจุดมุมเพื่อใช้เป็นจุดลักษณ์ การตรวจหาจุดลักษณ์นั้นทำเพื่อใช้เป็นข้อมูลในการหาความสัมพันธ์ระหว่างลำดับภาพ

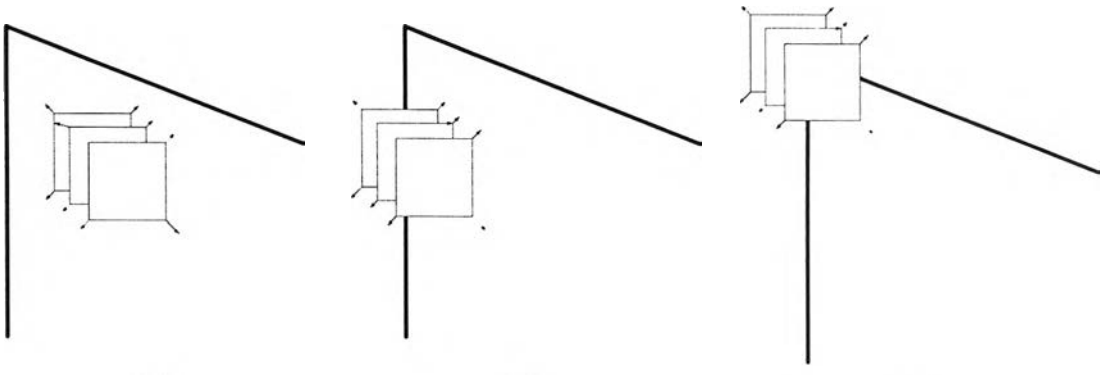
2.3.1.1 **การตรวจหาจุดมุมแบบแฮริส** (Harris corner detector)

การตรวจหาจุดมุมแบบแฮริส [5] ใช้วิธีการพิจารณาค่าการเปลี่ยนแปลงความเข้มแสงของแต่ละจุดภาพเทียบกับจุดภาพข้างเคียง ซึ่งค่าการเปลี่ยนแปลงความเข้มแสงสามารถแบ่งตามลักษณะการเปลี่ยนแปลงได้ 3 ลักษณะ

1.ค่าการเปลี่ยนแปลงความเข้มแสงมีขนาดเท่ากันในทุกๆ ด้านลักษณะนี้แสดงว่าภาพบริเวณนั้นเป็นบริเวณพื้นเรียบ แสดงในรูปที่ 2.4 (a) ซึ่งในรูปนี้บริเวณการค้นหาที่แสดงโดยกรอบสี่เหลี่ยมจัตุรัสทำการค้นหาอยู่ในบริเวณที่ค่าความเข้มแสงไม่มีการเปลี่ยนแปลงหรือเปลี่ยนแปลงน้อยมาก

2.ค่าการเปลี่ยนแปลงความเข้มแสงมีค่ามากเพียงแนวใดแนวหนึ่ง ลักษณะนี้แสดงว่าภาพบริเวณนั้นมีลักษณะเป็นเส้นขอบ แสดงในรูปที่ 2.4 (b) ซึ่งในรูปนี้บริเวณการค้นหาที่แสดงโดยกรอบสี่เหลี่ยมจัตุรัสทำการค้นหาอยู่ในบริเวณที่มีลักษณะเป็นเส้นขอบซึ่งมีค่าการเปลี่ยนแปลงความเข้มแสงในแนวขนานกับเส้นขอบน้อย แต่มีค่าการเปลี่ยนแปลงความเข้มแสงในแนวตั้งฉากกับเส้นขอบมาก

3.ค่าการเปลี่ยนแปลงความเข้มแสงมีค่ามากในทุกแนว ลักษณะนี้แสดงว่าภาพบริเวณนั้นมีลักษณะจุดหรือมุม แสดงในรูปที่ 2.4 (c) ซึ่งในรูปนี้บริเวณการค้นหาที่แสดงโดยกรอบสี่เหลี่ยมจัตุรัสทำการค้นหาอยู่ในบริเวณที่มีลักษณะเป็นจุดมุมซึ่งมีค่าการเปลี่ยนแปลงความเข้มแสงมากในหลายทิศทางเพราะจากรูปจุดมุมนี้เกิดจากเส้นขอบสองเส้นบรรจบกัน ดังนั้นในบริเวณค้นหาค่าการเปลี่ยนแปลงความเข้มแสงก็จะมีค่ามากในแนวตั้งฉากกับเส้นขอบทั้งสอง และในจุดที่เส้นมาบรรจบกันด้วย



(a) (b) (c)

รูปที่ 2.4 ลักษณะการเปลี่ยนแปลงของภาพ (a) พื้นผิวเรียบ (b) เส้นขอบ (c) จุดมุม

ดังนั้นการตรวจหาจุดมุมแบบแฮริส ก็คือการเลือกจุดการเปลี่ยนแปลงค่าความเข้มแสงมากในทุกทิศทางเพื่อเป็นจุดมุม

$$E(u,v)=\sum_{x,y} w(x,y)\left[I(x+u,y+v)-I(x,y)\right]^2 \tag{2.9}$$

$E(u,v)$ คือค่าการเปลี่ยนแปลงความเข้มแสงของภาพเมื่อเลื่อนตำแหน่งไป (u,v)

$I(x+u,y+v)$ คือความเข้มแสงของภาพเมื่อเลื่อนไป (u,v)

$I(x,y)$ คือความเข้มแสงของจุดภาพที่ต้องการวัด

$w(x,y)$ คือฟังก์ชันหน้าต่าง

การหาค่ามุมหาได้จากค่าการเปลี่ยนแปลงของ $E(u,v)$ ในทิศทางต่างๆ และเมื่อการเลื่อนที่ (u,v) มีขนาดเล็กสามารถประมาณค่า $E(u,v)$ ได้

$$E(u,v)\cong[u,v]M\begin{bmatrix}u\\v\end{bmatrix} \tag{2.10}$$

เมื่อ M คือ เมทริกซ์สมมาตรขนาด 2×2 ที่คำนวณจากค่าอนุพันธ์ของภาพ และเป็นเมทริกซ์ที่แสดงรูปร่างลักษณะข้อมูลของ $E(u,v)$

$$M=\sum_{x,y} w(x,y)\begin{bmatrix}I_x^2 & I_xI_y\\ I_xI_y & I_y^2\end{bmatrix} \tag{2.11}$$

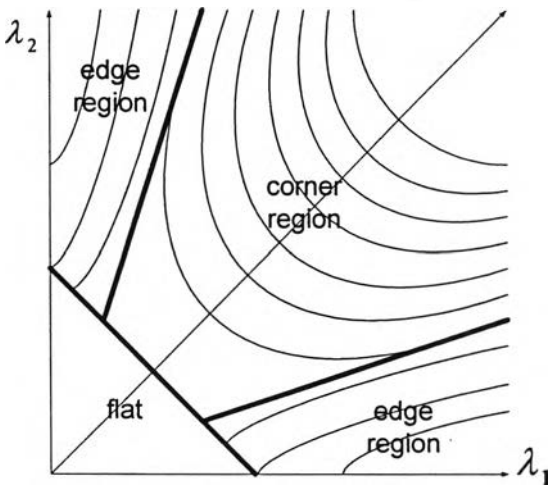
$$w(x,y)=e^{-\frac{(x^2+y^2)}{2\sigma^2}}$$

I_x คือค่าเกรเดียนต์ในแนวแกน $x\left(\frac{\partial I}{\partial x}\right)$ และ I_y คือค่าเกรเดียนต์ในแนวแกน $y\left(\frac{\partial I}{\partial y}\right)$ เนื่องจาก M มีคุณสมบัติเป็นเมตริกซ์สมมาตรและสามารถเขียนในรูปแบบของเมตริกซ์ทแยงมุม (Diagonal matrix) ที่มีค่า λ_1,λ_2 คือ ค่าเฉพาะ (eigenvalues) ของ M โดยค่าเฉพาะจะเป็นค่าที่แสดงการ

เปลี่ยนแปลงของข้อมูล $E(u,v)$ ว่ามีการเปลี่ยนแปลงในทิศทางใด และเปลี่ยนแปลงมากน้อยเพียงใด

$$M = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \tag{2.12}$$

จากค่าเมทริกซ์ M ที่ได้ สามารถนำไปหาค่าตอบสนองทางมุม (corner response) ซึ่งเป็นค่าที่แสดงการเปลี่ยนของข้อมูล $E(u,v)$



รูปที่ 2.5 การแบ่งลักษณะภาพโดยใช้ค่าเฉพาะ

จากรูปที่ 2.5 ในบริเวณที่ความเข้มแสงเท่ากันตลอดหรือมองเป็นพื้นผิวสีเดียวกัน (flat region) ค่า λ_1 และ λ_2 มีขนาดเล็ก ส่วนค่า E จะมีค่าการเปลี่ยนแปลงคงที่ในทุกทิศทาง ส่วนบริเวณที่เป็นเส้นขอบ (edge region) ค่า λ_1 และ λ_2 จะมีค่าต่างกันมากคือ $\lambda_1 \ll \lambda_2$ หรือ $\lambda_1 \gg \lambda_2$ ส่วนค่า E จะมีค่าการเปลี่ยนแปลงมากในทิศทางใดทิศหนึ่ง ส่วนบริเวณที่เป็นจุดมุม (corner region) ค่า λ_1 และ λ_2 จะมีค่ามากคือ $\lambda_1 \sim \lambda_2$ ส่วนค่า E จะมีค่าการเปลี่ยนแปลงมากในทุกทิศทาง จึงต้องหาค่าการตอบสนองทางมุมเพื่อใช้วัดค่าการเปลี่ยนแปลง

$$\begin{aligned} R &= \det M - k\left(trace(M)\right)^2 \\ \det M &= \lambda_1\lambda_2 = I_x^2 I_y^2 - \left(I_x I_y\right)^2 \\ trace(M) &= \lambda_1 + \lambda_2 = I_x^2 + I_y^2 \end{aligned} \tag{2.13}$$

R คือค่าตอบสนองทางมุม (corner response) ส่วน k คือค่าคงที่เชิงประสบการณ์ (empirical constant) มีค่าระหว่าง 0.04-0.06 ในการเลือกค่าจุดมุมจะพิจารณาจากค่า R ซึ่งค่า R จะขึ้นอยู่กับค่า ค่าเฉพาะ (eigenvalues) ของ M ดังนี้

1. ค่า $R > 0$ คือจุดมุม
2. ค่า $R < 0$ คือเส้นขอบ
3. ค่า $|R| \approx 0$ คือผิวเรียบ

ดังนั้นการเลือกมุมที่จะใช้เป็นจุดลักษณ์เลือกจากจุดที่มีค่าตอบสนองทางมุมมากที่สุด ซึ่งก็ปรากฏได้หลายจุดจึงอาจใช้วิธีกำหนดให้มีค่ามากกว่าค่าขีดแบ่ง (threshold) ที่ตั้งไว้ เพื่อให้ได้จุดมุมที่ชัดเจน

2.3.2 การติดตามจุดลักษณ์ (Feature tracking)

การติดตามจุดลักษณ์ คือการหาตำแหน่งของจุดลักษณ์ที่เลื่อนไปในลำดับภาพถัดไป ซึ่งก็คือการหาการเลื่อนที่หรือหาค่าเวกเตอร์การเคลื่อนที่นั่นเอง วิธีหาค่าเวกเตอร์การเคลื่อนที่นั้นแบ่งเป็นการหาการเคลื่อนที่เชิงเวลา และการหาการเคลื่อนที่เชิงความถี่ ซึ่งในวิทยานิพนธ์ฉบับนี้เลือกใช้วิธีการเข้าคู่บล็อกในเชิงเวลาเพราะมีความซับซ้อนต่ำและให้ความถูกต้องในระดับที่ยอมรับได้

ระบบการประมาณการเคลื่อนที่เชิงเวลานั้นมีหลายวิธี

วิธีการเข้าคู่บล็อกเหมาะสมแบบค้นเต็มที่ (Full-search block matching algorithm)

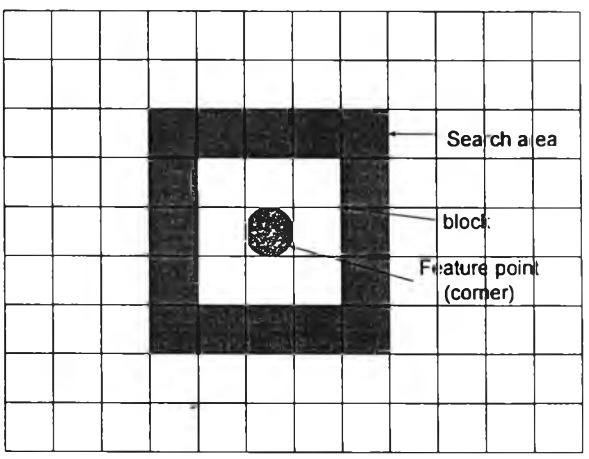
โดยวิธีการเข้าคู่บล็อกเหมาะสมแบบค้นเต็มที่ จะทำการแบ่งรูปเป็นบล็อกขนาดเท่าๆกัน แล้วสมมติว่าแต่ละบล็อกมีเวกเตอร์การเคลื่อนที่ร่วมกัน กำหนดค่าระยะการค้นหาและทำการค้นหาในระยะการค้นโดยเปรียบเทียบเฟรมภาพปัจจุบันกับเฟรมภาพอ้างอิง แล้วเลือกบล็อกที่เหมาะสมโดยใช้เงื่อนไขค่าความแตกต่างสัมบูรณ์เฉลี่ย และค่าผิดพลาดกำลังสองเฉลี่ย แต่วิธีการเข้าคู่บล็อกเหมาะสมแบบค้นเต็มที่นั้น แม้จะมีความผิดพลาดต่ำสุด แต่ต้องใช้การคำนวณมาก ส่งผลให้เกิดการประวิงเวลา และยังต้องใช้อุปกรณ์ที่มีความซับซ้อนสูงจึงไม่เหมาะที่จะนำไปใช้กับกล้องวีดีโอ [4]

วิธีการเข้าคู่จุดแทน

วิธีการเข้าคู่จุดแทนจะแบ่งเฟรมภาพออกเป็นส่วนการหาค่าเวกเตอร์การเคลื่อนที่สี่ส่วน กำหนดจุดแทนในแต่ละส่วนการหาค่าเวกเตอร์การเคลื่อนที่ ใช้จุดแทนนี้ในการหาค่าเวกเตอร์การเคลื่อนที่ของแต่ละส่วนโดยใช้การหาค่าสหสัมพันธ์(Correlation) เลือกตำแหน่งที่มีค่าสหสัมพันธ์

ตามที่กำหนดเป็นตำแหน่งของเวกเตอร์การเคลื่อนที่ ค่าเวกเตอร์การเคลื่อนที่รวมจะได้จากการหาผลรวมของเวกเตอร์การเคลื่อนที่ของแต่ละส่วน แต่วิธีการนี้ มีปัญหาที่ระบบมีความไวต่อเงื่อนไขความผิดปกติ เช่น มีวัตถุเคลื่อนที่ผ่าน การสายกล้องแนวราบและการหมุนกล้อง [6]

การติดตามจุดลักษณ์ในที่นี้มีวิธีการดังนี้ คือเมื่อได้จุดลักษณ์ซึ่งก็คือจุดมุมในภาพอ้างอิงที่ก็จะทราบค่าพิกัด และความเข้มแสงของจุดลักษณ์ โดยการติดตามจุดลักษณ์ที่เลื่อนที่ไปทำได้โดยกำหนดหน้าต่างขนาดเล็กครอบบนจุดพิกัดจุดลักษณ์ของภาพอ้างอิงนั้นเรียกว่า บล็อก แล้วใช้บล็อกนี้ไปทำการหาค่าสหสัมพันธ์ (Correlation) ซึ่งเป็นวิธีการหาความสัมพันธ์ระหว่างตัวแปรเทียบกับภาพเป้าหมาย โดยจะค้นหาในขอบเขตที่กำหนดเรียกว่า พื้นที่ค้นหา (search area) ของภาพเป้าหมายดังรูปที่ 2.6 ซึ่งเป็นขอบเขตพื้นที่ที่มีความเป็นไปได้ที่จะเกิดเวกเตอร์การเคลื่อนที่สูงสุด แล้วเลือกตำแหน่งที่มีค่าสหสัมพันธ์มากที่สุดเป็นค่าพิกัดจุดสักษณ์ที่เปลี่ยนไปในภาพเป้าหมาย



รูปที่ 2.6 ลักษณะการสร้างบล็อกครอบจุดลักษณ์ และพื้นที่ค้นหา

การหาค่าสหสัมพันธ์ (Correlation) เป็นการหาความสัมพันธ์ระหว่างข้อมูล โดยพิจารณาจากค่าสัมประสิทธิ์สหสัมพันธ์ (coefficient of correlation) ในที่นี้คือค่า r ซึ่งค่า r ที่มากจะแสดงให้เห็นว่าข้อมูลที่นำมาเปรียบเทียบกันนั้นมีความสัมพันธ์กันมาก ส่วนค่า r ที่น้อยจะแสดงให้เห็นว่าข้อมูลที่นำมาเปรียบเทียบกันนั้นไม่ค่อยมีความสัมพันธ์กัน

$$r = \frac{\sum_x \sum_y \left(I_{A_{xy}} - \bar{I}_{A_{xy}}\right)\left(I_{B_{xy}} - \bar{I}_{B_{xy}}\right)}{\sqrt{\left(\sum_x \sum_y \left(I_{A_{xy}} - \bar{I}_{A_{xy}}\right)^2\right)\left(\sum_x \sum_y \left(I_{B_{xy}} - \bar{I}_{B_{xy}}\right)^2\right)}} \tag{2.14}$$

r คือสัมประสิทธิ์สหสัมพันธ์

$I_{A_{xy}}, I_{B_{xy}}$ คือค่าความเข้มแสงของภาพ A และภาพ B ที่จุด (x, y)

$\overline{I}_{A_{xy}}, \overline{I}_{B_{xy}}$ คือค่าเฉลี่ยของความเข้มแสงในหน้าต่างของภาพ A และภาพ B

x และ y คือตำแหน่งพิกัดในส่วนข้อมูลที่นำมาหาค่าสหสัมพันธ์

จากสมการที่ 2.14 จะเลือกจุดที่มีค่า r มากที่สุดในบริเวณพื้นที่ค้นหาเป็นจุดลักษณ์รูปเป้าหมาย ซึ่งก็ทราบค่าพิกัดจุดลักษณ์ ก็นำค่าจุดพิกัดจากรูปอ้างอิงและจุดพิกัดที่หาได้จากรูปเป้าหมายนั้นไปทำการแปลงสัมพรรคเพื่อดึงภาพเป้าหมายที่หมุนไปให้กลับมาอยู่ตำแหน่งเดิมกับภาพอ้างอิง

2.4. การบิดภาพ(Image Warping)

การบิดภาพ(Image Warping) เป็นขั้นตอนในการสร้างภาพกลับขึ้นมาใหม่คือบิดภาพที่หมุนไปให้กลับมาที่ตำแหน่งเดิม โดยอาศัยค่าการแปลงสัมพรรค (a_0 a_1 a_2 b_0 b_1 และ b_2) การหาค่าการแปลงสัมพรรคนั้นต้องทราบการเปลี่ยนแปลงของภาพในที่นี้คือทราบพิกัดของภาพที่เปลี่ยนไป นำค่าการเลื่อนที่จากสมการที่ 2.8 เขียนได้ดังนี้

$$x' = x + d_x(x, y)$$
$$y' = y + d_y(x, y)$$
$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} + \begin{bmatrix} a_1 & a_2 & a_0 \\ b_1 & b_2 & b_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \quad (2.15)$$

x' และ y' คือค่าพิกัดของภาพหลังจากเกิดการเคลื่อนที่

x และ y คือค่าพิกัดของภาพก่อนเกิดการเคลื่อนที่ (ภาพอ้างอิง)

นำจุดพิกัดที่ได้จากหาค่าสหสัมพันธ์มาเขียนสมการเพื่อหาค่าการแปลงสัมพรรคได้สมการที่ 2.16 และสมการที่ 2.17

$$\begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} 1 & x_1 & y_1 \\ 1 & x_2 & y_2 \\ 1 & x_3 & y_3 \end{bmatrix}^{-1} \begin{bmatrix} x_1' - x_1 \\ x_2' - x_2 \\ x_3' - x_3 \end{bmatrix} \quad (2.16)$$

$$\begin{bmatrix} b_0 \\ b_1 \\ b_2 \end{bmatrix} = \begin{bmatrix} 1 & x_1 & y_1 \\ 1 & x_2 & y_2 \\ 1 & x_3 & y_3 \end{bmatrix}^{-1} \begin{bmatrix} y_1' - y_1 \\ y_2' - y_2 \\ y_3' - y_3 \end{bmatrix} \quad (2.17)$$

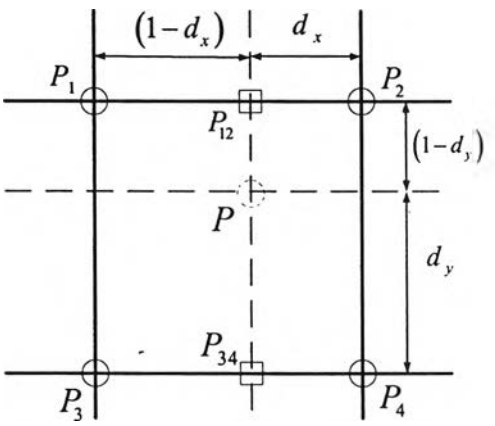
โดยจำนวนจุดที่นำมาใช้อย่างน้อยที่สุดคือ 3 จุด เมื่อได้ค่าการแปลงสัมพรรคแล้วก็นำมาใช้ในการบิดภาพกลับซึ่งการบิดภาพกลับในที่นี้หมายถึงการดึงภาพที่เลื่อนหรือเปลี่ยนตำแหน่งไปให้กลับมาที่ตำแหน่งเดิม

$$\begin{bmatrix} x \\ y \end{bmatrix} = \left(\begin{bmatrix} x' \\ y' \end{bmatrix} - \begin{bmatrix} a_0 \\ b_0 \end{bmatrix} \right) (I + A)^{-1}$$
$$A = \begin{bmatrix} a_1 & a_2 \\ b_1 & b_2 \end{bmatrix} \quad (2.18)$$

จากสมการที่ 2.18 ค่า (x, y) คือพิกัดของภาพอ้างอิง (x', y') คือพิกัดภาพที่เลื่อนที่ไป I คือเมทริกซ์เอกลักษณ์ (Identity matrix)

2.4.1 การประมาณค่าในช่วงเชิงเส้นคู่ (Bilinear interpolation)

ในการคำนวณพิกัด (x, y) จากสมการที่ 2.18 ค่าที่ได้อาจเป็นค่าทศนิยมซึ่งไม่สามารถใช้ดึงข้อมูลในระบบพิกัดของภาพได้โดยตรงจึงต้องมีวิธีการประมาณค่าข้อมูลในที่นี้ใช้วิธีการประมาณค่า 2 มิติ แบบการประมาณค่าในช่วงเชิงเส้นคู่ (Bilinear interpolation)



รูปที่ 2.7 การประมาณค่าในช่วงเชิงเส้นคู่ (Bilinear interpolation)

จากรูปที่ 2.7 จุด P เป็นพิกัด (x, y) ที่คำนวณได้ ส่วน P_1, P_2, P_3 และ P_4 เป็นจุดพิกัดข้างเคียงที่ใช้ช่วยประมาณค่า P สามารถคำนวณค่า P ได้ดังสมการที่ 2.19

$$\begin{aligned} P_{12} &= d_x P_1 + (1 - d_x) P_2 \\ P_{34} &= d_x P_3 + (1 - d_x) P_4 \\ P &= d_y P_{12} + (1 - d_y) P_{34} \\ &= d_x d_y P_1 + (1 - d_x) d_y P_2 + d_x (1 - d_y) P_3 + (1 - d_x)(1 - d_y) P_4 \end{aligned} \quad (2.19)$$